Network Working Group
Request for Comments: 3643
Category: Standards Track

R. Weber
Brocade
M. Rajagopal
Broadcom Corporation
F. Travostino
Nortel Networks
M. O'Donnell
McDATA
C. Monia
Nishan Systems
M. Merhar
Sun Microsystems

December 2003

Fibre Channel (FC) Frame Encapsulation

Status of this Memo

Copyright Notice

Abstract

This document describes the common Fibre Channel (FC) frame encapsulation format and a procedure for the measurement and calculation of frame transit time through the IP network. This specification is intended for use by any IETF protocol that encapsulates FC frames.

Table Of Contents

1. Scope

This document describes common mechanisms for the transport of Fibre Channel frames over an IP network, including the encapsulation format and a mechanism for enforcing the Fibre Channel frame lifetime limits.

Warning to Readers Familiar With Fibre Channel: Both Fibre Channel and IETF standards use the same byte transmission order. However, the bit and byte numbering is different. See Appendix A for guidance.

The organization responsible for the Fibre Channel standards (INCITS Technical Committee T11) has determined that some functions and modes of operation are not interoperable to the degree required by the IETF (see FC-MI [8]). This document includes applicable T11 interoperability determinations in the form of restrictions on the use of this encapsulation mechanism.

Use of these mechanisms in an encapsulating protocol requires an additional document to specify the encapsulating protocol specific functionality and appropriate security considerations. Because security considerations for this encapsulation depend on how it is used by encapsulating protocols, they are taken up in encapsulating protocol specific documents.

Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119 [2].

2. Encapsulation Concepts

The smallest unit of data transmission and routing in Fibre Channel (FC) is the frame. FC frames include a Start Of Frame (SOF), End Of Frame (EOF), and the contents of the Fibre Channel frame. The Fibre Channel frame includes a Cyclic Redundancy Check (CRC) code that provides error detection for the contents of the frame. FC frames are variable length. To facilitate transporting FC frames over an IP based transport such as TCP the native FC frame needs to be contained in (encapsulated in) a slightly larger structure as shown in Figure 1.

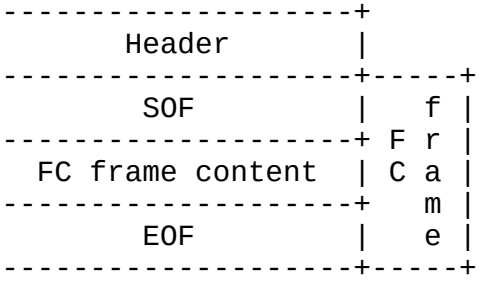

Figure 1 - FC frame Encapsulation

The format and content of an FC frame are described in the FC standards (e.g., FC-FS [3], FC-SW-2 [4], and FC-PI [5]). Of importance to this encapsulation is the FC requirement that all frames SHALL contain a CRC for detection of transmission errors.

3. The FC Encapsulation Header

3.1. FC Encapsulation Header Format

Figure 2 shows the format of the required FC Encapsulation Header.

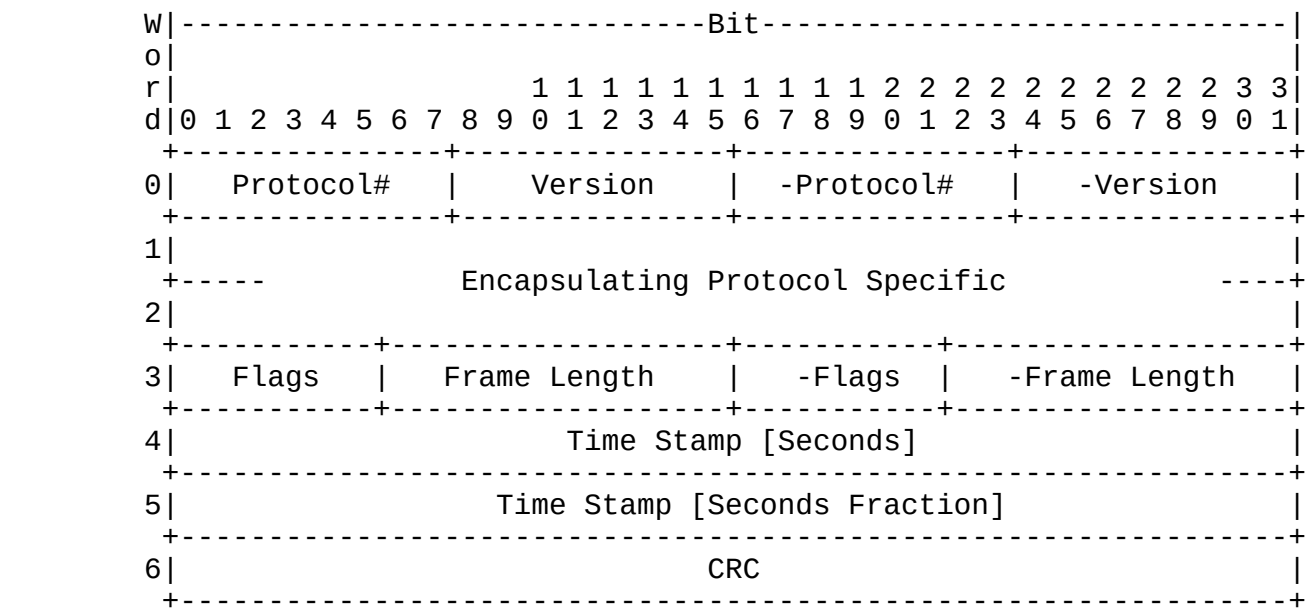

Figure 2 - FC Encapsulation Header Format

The fields in the FC Encapsulation Header are defined as follows.

Protocol#: The Protocol# field SHALL contain a number that indicates which encapsulating protocol is employing the FC Encapsulation. The values in the Protocol# field are assigned by IANA (see Appendix C).

Version: The Version field SHALL contain 0x01 to indicate that this version of the FC Encapsulation is being used. All other values are reserved for future versions of the FC Encapsulation.

-Protocol#: The -Protocol# field SHALL contain the one’s complement of the contents of the Protocol# field. FC Encapsulation receivers SHOULD either validate the CRC or compare the Protocol# and - Protocol# fields to verify that an FC Encapsulation Header is being processed according to a policy defined by the encapsulating protocol.

-Version: The -Version field SHALL contain the one's complement of the contents of the Version field. FC Encapsulation receivers SHOULD either validate the CRC or compare the Version and -Version fields to verify that an FC Encapsulation Header is being processed according to a policy defined by the encapsulating protocol.

Encapsulating Protocol Specific: The usage of these words differs based on the contents of the Protocol# field, i.e., the usage of these words is defined by the encapsulating protocol that is employing this encapsulation.

Flags: The Flags bits provide information about the usage of the FC Encapsulation Header as shown in Figure 3.

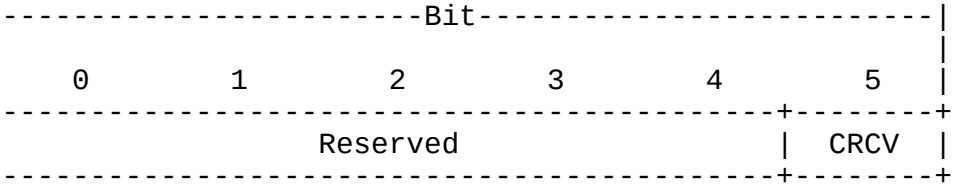

Figure 3 - Flags Field Format

Reserved Flags bits: These bits are reserved for use by future versions of the FC Encapsulation and SHALL be set to zero on send. Encapsulating protocols employing the encapsulation described in this specification MAY require checking for zero on receive, however doing so has the potential to create incompatibilities with future versions of this encapsulation. Changes in the usage of the Reserved Flags bits MUST be identified by changes in the contents of the Version field. Encapsulating protocols employing the encapsulation described in this specification MUST NOT make use of the Reserved Flags bits in any fashion other than that described in this specification.

CRCV (CRC Valid Flag): A CRCV bit value of one indicates that the contents of the CRC field are valid. A CRCV bit value of zero indicates that the contents of the CRC field are invalid. The value of the CRCV bit SHALL be constant for all FC Encapsulation Headers sent on a given connection.

Frame Length: The Frame Length field contains the length of the entire FC Encapsulated frame including the FC Encapsulation Header and the FC frame (including SOF and EOF words). This length is based on a unit of 32-bit words. All FC frames are 32-bit-word-aligned and the FC Encapsulation Header is always word-aligned; therefore a32-bit word length is acceptable.

-Flags: The -Flags field SHALL contain the one's complement of the contents of the Flags field. FC Encapsulation receivers SHOULD either validate the CRC or compare the Flags and -Flags fields to verify that an FC Encapsulation Header is being processed according to a policy defined by the encapsulating protocol.

-Frame Length: The -Frame Length field SHALL contain the one's complement of the contents of the Frame Length field. FC Encapsulation receivers SHOULD either validate the CRC or compare the Frame Length and -Frame Length fields to verify that an FC Encapsulation Header is being processed according to a policy defined by the encapsulating protocol.

Time Stamp [Seconds]: The Time Stamp [Seconds] field contains zero or the number of seconds since 0 hour on 1 January 1900 at the time the FC Encapsulated frame is place in the outgoing data stream.

Time Stamp [Seconds Fraction]: The Time Stamp [Second Fraction] field contains the fraction of the second at the time the FC Encapsulated frame is place in the outgoing data stream. Non-significant low order bits may be set to zero. Table 1 shows some example Time Stamp [Seconds Fraction] values.

Second	Time Stamp [Seconds Fraction]
n.50000...	0x80000000
n.25000...	0x40000000
n.12500...	0x20000000

Table 1 Example Time Stamp [Seconds Fraction] values

Note that, since some time in 1968 (second 2,147,483,648) the most significant bit (bit 0 of Time Stamp [Seconds]) has been set and that the field will overflow some time in 2036 (second 4,294,967,296). Should FCIP be in use in 2036, some external means will be necessary to qualify time relative to 1900 and time relative to 2036 (and other multiples of 136 years). There will exist a 200-picosecond interval, henceforth ignored, every 136 years when the 64-bit field will be 0, which by convention is interpreted as an invalid or unavailable timestamp.

The Time Stamp [Seconds] and Time Stamp [Seconds Fraction] words follow the in time format described in Simple Network Time Protocol (SNTP) Version 4 [9]. The contents of the Time Stamp [Seconds] and Time Stamp [Seconds Fraction] words SHALL be set as described in section 4.

CRC: When the CRCV Flag bit is zero, the CRC field SHALL contain zero. When the CRCV Flag bit is one, the CRC field SHALL contain a CRC for words 0 to 5 of the FC Encapsulation Header computed using the equations, polynomial, initial value, and bit order defined for Fibre Channel in FC-FS [3]. Using this algorithm, the bit order of the resulting CRC corresponds to that of FC-1 layer. The CRC transmitted over the IP network shall correspond to the equivalent value converted to FC-2 format as specified in FC-FS.

3.2. FC Encapsulation Header Validation

Two mechanisms are provided for validating an FC Encapsulation Header:

- Redundancy based
- CRC based

The two mechanisms address the needs of two different design and operating environments.

3.2.1. Redundancy Based FC Encapsulation Header Validation

Redundancy based validation of an FC Encapsulation Header relies on duplicated and one’s complemented fields in the header.

Encapsulating protocols that use redundancy based validation SHOULD define how receiving devices use one’s complement fields to verify header validity.

Header validation based on redundancy is a stepwise process in that the first word is validated, then the second, then the third and so on. A decision that a candidate header is not valid may be reached before the complete header is available.

3.2.2. CRC Based FC Encapsulation Header Validation

CRC based validation of an FC Encapsulation Header relies on a CRC located in the last word of the header.

Header validation based on the CRC defined in section 3.1 requires computing the CRC for all bytes preceding the CRC word, and comparing the results to the CRC word’s contents.

4. Measuring Fibre Channel Frame Transit Time

To comply with FC-FS [3], an FC Fabric must specify and limit the lifetime of a frame. In an FC Fabric comprised of IP-connected elements, one component of the frame's lifetime is the time required to traverse the connection. To ensure that the total frame lifetime remains within the limits required by the FC Fabric, the encapsulation described in this specification contains provisions for recording the departure time of an encapsulated frame injected into a connection. If the encapsulated frame originator and recipient have access to aligned and synchronized time bases, the transit time through the IP network can then be computed.

When originating an encapsulated frame, an entity that does not support transit time calculation SHALL always set the Time Stamp [Seconds] and Time Stamp [Seconds Fraction] fields to zero. When receiving an encapsulated frame, an entity that does not support transit time calculation SHALL ignore the contents of the Time Stamp words.

The encapsulating protocol SHALL specify whether or not implementation support is required. The encapsulating protocol SHALL specify those conditions under which a received encapsulated frame MUST have its transit time checked before forwarding.

Encapsulating and de-encapsulating entities that support this feature MUST have access to:

a) An internal time base having the stability and resolution necessary to comply with the requirements of the encapsulating protocol specification; and

b) A time base that is synchronized and aligned with the time base of other entities to which encapsulated frames may be sent or received. The encapsulating protocol specification MUST describe the synchronization and alignment procedure.

With respect to its ability to measure and set transit time for encapsulated frames exchanged with another device, an entity is either in the Synchronized or Unsynchronized state. An entity is in the Unsynchronized state upon power-up and transitions to the Synchronized state once it has aligned its time base in accordance with the applicable encapsulating protocol specification.

An entity MUST return to the Unsynchronized state if it is unable to maintain synchronization of its time base as required by the encapsulating protocol specification.

The policy for forwarding frames while in the Unsynchronized state SHALL be defined by the encapsulating protocol specification.

If processing frames in the Unsynchronized state is permitted by the encapsulating protocol specification, the entity SHALL:

a) When de-encapsulating a frame, ignore the Time Stamp words. For example, if a calculated transit time exceeds a value that could cause the frame to violate FC maximum time in transit limits, the encapsulating protocol may specify that the frame is to be discarded; and

b) When encapsulating a frame set the Time Stamp [Seconds] and Time Stamp [Seconds Fraction] words to zero. For example, an encapsulating protocol may specify that frames for which transit time cannot be determined are never to be forwarded over FC.

When encapsulating a frame, an entity in the Synchronized state SHALL record the value of the time base in the Time Stamp [Seconds] and Time Stamp [Seconds Fraction] words in the encapsulation header.

When de-encapsulating a frame, an entity in the Synchronized state SHALL:

a) Test the Time Stamp words to determine if they contain a time as specified in [9]. If the time stamp is valid, the receiving entity SHALL compute the transit time by calculating the difference between its time base and the departure time recorded in the frame header. The receiving entity SHALL process the calculated transit time and the de-encapsulated frame in accordance with the applicable encapsulating protocol specification; or

b) If both Time Stamp words have a value of zero, the receiving entity SHALL de-encapsulate the frame without computing the transit time. The disposition of the frame and any other actions by the recipient SHALL be defined by the encapsulating protocol specification.

Note: For most purposes, communication between entities is possible only while in the Synchronized state.

5. The FC Frame

5.1. FC Frame Content

NOTE: All uses of the words "character" or "characters" in this section refer to 8bit/10bit link encoding wherein each 8 bit "character" within a link frame is encoded as a 10 bit "character" for link transmission. These words do not refer to ASCII, Unicode, or any other form of text characters, although octets from such characters will occur as 8 bit "characters" for this encoding. This usage is employed here for consistency with the ANSI T11 standards that specify Fibre Channel.

Figure 4 shows the structure of a general FC-2 frame format.

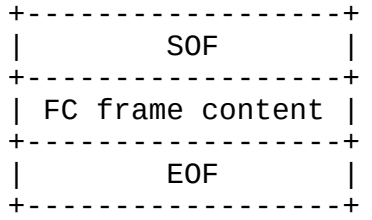

Figure 4 - General FC-2 Frame Format

As shown in Figure 4, the FC frame content is defined as the data between the EOF and SOF delimiters (including the FC CRC) after conversion from FC-1 to FC-2 format as specified by FC-FS [3].

When Fibre Channel devices convert the FC frame content to the FC-0 physical transport, an encoding is applied to the FC frame content (e.g., 8b/10b encoding like that used in Gigbit Ethernet) for reasons that include redundancy and low level timing synchronization between sender and receiver. With the exceptions of SOF and EOF [3] all discussion of FC frame content in this document is at the 8-bit byte level, prior to the application of any such encoding.

The 8-bit bytes in the FC frame content can be translated directly for transmission over an IP Network. However, the FC SOF and EOF employ special 10b characters that have no 8b equivalents. Therefore, special byte placement and 8-bit character encodings are required to represent SOF and EOF.

5.2. Bit and Byte Ordering

The Encapsulation Header, SOF, FC frame content (see section 5.1), and EOF are mapped to TCP using the big endian byte ordering, which corresponds to the standard network byte order or canonical form [7].

5.3. FC SOF and EOF

As described in section 5.1, representation of FC SOF and EOF in an IP Network byte stream requires special formatting and 8-bit code definitions. Therefore, the encapsulated FC frame SHALL have the format shown in Figure 5. The redundancy of the SOF/EOF representation in the encapsulation format results from concerns that the information be protected from transmission errors.

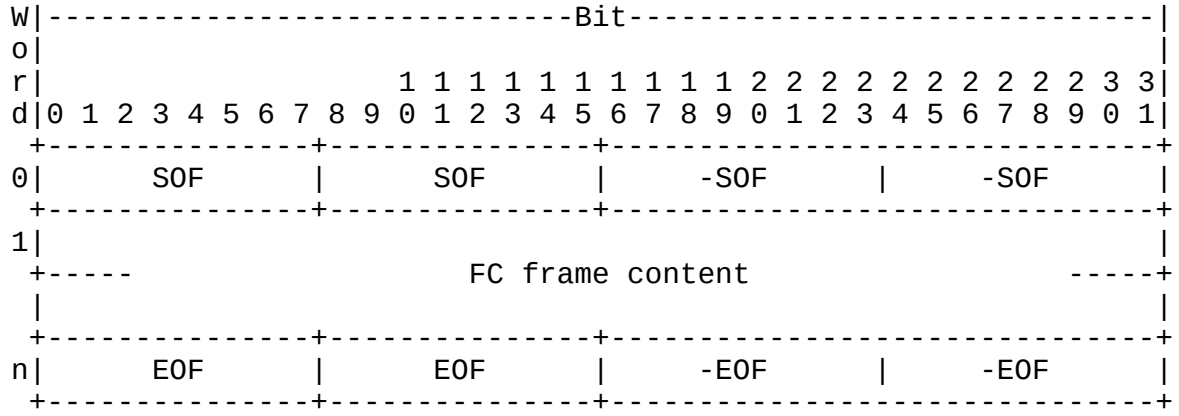

Figure 5 - FC Frame Encapsulation Format

Note: The number of 8-bit bytes in the FC frame content is always a multiple of four.

SOF: The SOF fields contain the encoded SOF value selected from table 2.

FC SOF	SOF Code	Class
SOFf	0x28	F
SOFi2	0x2D	2
SOFn2	0x35	2
SOFi3	0x2E	3
SOFn3	0x36	3
SOFi4	0x29	4
SOFn4	0x31	4
SOFc4	0x39	4

Table 2 Translation of FC SOF values to SOF field contents

-SOF: The -SOF fields contain the one’s complement of the value in the SOF fields. Encapsulation receivers SHOULD validate the SOF field according to a policy defined by the encapsulating protocol.

EOF: The EOF fields contain the encoded EOF value selected from table 3.

FC EOF	EOF Code	Class	FC EOF	EOF Code	Class
EOFn	0x41	2,3,4,F	EOFdt	0x46	4
EOFt	0x42	2,3,4,F	EOFdti	0x4E	4
EOFni	0x49	2,3,4,F	EOFrt	0x44	4
EOFa	0x50	2,3,4,F	EOFrti	0x4F	4

Table 3 Translation of FC EOF values to EOF field contents

-EOF: The -EOF fields contain the one's complement of the value in the EOF fields. Encapsulation receivers SHOULD validate the EOF field according to a policy defined by the encapsulating protocol.

Note: FC-BB-2 [6] lists SOF and EOF codes not shown in table 2 and table 3 (e.g., SOFi1 and SOFn1). However, FC-MI [8] identifies these codes as not interoperable, so they are not listed in this specification.

6. Security Considerations

This document describes the encapsulation format only. Actual use of this format in a encapsulating protocol requires an additional document to specify the encapsulating protocol functionality and appropriate security considerations. Because security considerations for this encapsulation depend on how it is used by encapsulating protocols, they SHALL be described in encapsulating protocol specific documents.

7. References

7.1. Normative References

[1] Bradner, S., "The Internet Standards Process -- Revision 3", BCP 9, RFC 2026, October 1996.

[2] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[3] Fibre Channel Framing and Signaling (FC-FS), ANSI INCITS.373:2003, October 27, 2003. Note: Published T11 standards are available from the INCITS online store http://www.incits.org, or the ANSI online store, http://www.ansi.org.

[4] Fibre Channel Switch Fabric -2 (FC-SW-2), ANSI NCITS.355:2001, December 12, 2002. Note: Published T11 standards are available from the INCITS online store http://www.incits.org, or the ANSI online store, http://www.ansi.org.

[5] Fibre Channel Physical Interfaces (FC-PI), ANSI NCITS.352:2002, December 1, 2002. Note: Published T11 standards are available from the INCITS online store http://www.incits.org, or the ANSI online store, http://www.ansi.org.

[6] Fibre Channel Backbone -2 (FC-BB-2), ANSI INCITS.372:2003, July 25, 2003. Note: Published T11 standards are available from the INCITS online store http://www.incits.org, or the ANSI online store, http://www.ansi.org.

[7] Narten, T. and C. Burton, "A Caution on The Canonical Ordering of Link-Layer Addresses", RFC 2469, December 1998.

7.2. Informative References

[8] Fibre Channel Methodologies for Interconnects (FC-MI), ANSI INCITS/TR-30:2002, November 1, 2002. Note: Published T11 standards are available from the INCITS online store http://www.incits.org, or the ANSI online store, http://www.ansi.org.

[9] Mills, D., "Simple Network Time Protocol (SNTP) Version 4 for IPv4, IPv6 and OSI", RFC 2030, October 1996.

[10] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.

[11] Rajagopal, M., Rodriguez, E., Weber, R., "Fibre Channel Over TCP/IP (FCIP)", Work in Progress.

[12] Monia, C., et. al., "iFCP - A Protocol for Internet Fibre Channel Storage Networking", Work in Progress.

8. Acknowledgements

The authors express their appreciation to Mr. Vi Chau (vchau1@cox.net) for his contributions to the design team that developed this document. Mr. Chau is no longer working in this technology.

The authors are also grateful to Dr. David Black, Mr. Mallikarjun Chadalapaka, and Mr. Robert Elliott for their reviews of this specification.

Appendix A - Fibre Channel Bit and Byte Numbering Guidance

Both Fibre Channel and IETF standards use the same byte transmission order. However, the bit and byte numbering is different.

Fibre Channel bit and byte numbering can be observed if the data structure heading shown in Figure 6, is cut and pasted at the top of Figure 2 and Figure 5.

```
W|-------------------------------Bit-------------------------------|
o|                                                                 |
r|3 3 2 2 2 2 2 2 2 2 2 1 1 1 1 1 1 1 1 1                          |
d|1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0 9 8 7 6 5 4 3 2 1 0|
```

Figure 6 - Fibre Channel Data Structure Bit and Byte Numbering

Fibre Channel bit numbering for the Flags field can be observed if the data structure heading shown in Figure 7, is cut and pasted at the top of Figure 3.

```
|------------------------Bit-------------------------|
|                                                     |
|   31       30       29       28       27       26   |
```

Figure 7 - Fibre Channel Flags Bit Numbering

Appendix B - Encapsulating Protocol Requirements

This appendix lists the requirements placed on the encapsulating protocols that employ this encapsulation. The requirements listed here are suggested or described elsewhere in this document, but their collection in this appendix serves to assist encapsulating protocol authors in meeting all obligations placed upon them.

Encapsulating Protocol Specific Data

Encapsulating protocols employing this encapsulation SHALL:

- specify the IANA assigned number used in the Protocol# field
- specify the contents of the Encapsulating Protocol Specific field

Encapsulating protocols employing this encapsulation SHALL define the procedures and policies necessary for verifying that an FC Encapsulation Header is being processed.

Encapsulating protocols employing this encapsulation SHALL define the procedures and policies necessary for the detection of over age frames. The items to be specified and the choices available to an encapsulating protocol specification are as follows:

a) The encapsulating protocol requirements for measuring transit times. The encapsulating protocol MAY allow implementation of transit time measurement to be optional.

b) The requirements or guidelines for stability and resolution of the entity's time base.

c) The procedure for synchronizing an entity's time base, including the criteria for entering the Synchronized and Unsynchronized states.

d) The forwarding (or lack of forwarding) of frame traffic while in the Unsynchronized state.

   The specification MAY allow an entity in the Unsynchronized state to continue processing frame traffic.

e) The procedure to be followed when frames are received that do not have a valid time stamp.

   The specification MAY allow such frames to be accepted by the entity.

f) Requirements for setting and testing the transit time limit and the procedure to be followed when a received frame is discarded due to its transit time exceeding the limit.

Appendix C - IANA Considerations

The Protocol# (Protocol Number) field is an identifier number used to distinguish between the encapsulating protocols that employ this FC frame encapsulation. Values used in the Protocol# field are to be assigned from a new, separate registry that is maintained by IANA.

All values in the Protocol# field are to be registered with and assigned by IANA with the following exceptions.

- Protocol# value 0 should not be assigned until after all other values have been assigned.

- Protocol# values 240-255 inclusive must be set aside for private use amongst cooperating systems.

Following the policies outlined in [10], Protocol# values not listed above are to be assigned only for Standards Track RFCs approved by the IESG.

In addition to creating the FC Frame Encapsulation Protocol Number Registry, the standards action of this RFC allocates the following two values from the registry:

- Protocol# value 1 assigned to the FCIP (Fibre Channel Over TCP/IP) encapsulating protocol [11].

- Protocol# value 2 assigned to the iFCP (A Protocol for Internet Fibre Channel Storage Networking) encapsulating protocol [12].

Appendix D - Intellectual Property Rights Statement

Authors' Addresses

Ralph Weber
ENDL Texas
representing Brocade Comm.
Suite 102 PMB 178
18484 Preston Road
Dallas, TX 75252
USA

Phone: +1 214 912 1373
EMail: roweber@ieee.org

Murali Rajagopal
Broadcom
16215 Alton Parkway
PO Box 57013
Irvine, CA 92619
USA

Phone: +1 949 450 8700
EMail: muralir@broadcom.com

Franco Travostino
Technology Center
Nortel Networks, Inc.
600 Technology Park
Billerica, MA 01821
USA

Phone: +1 978 288 7708
EMail: travos@nortelnetworks.com

Michael E. O'Donnell
McDATA Corporation
4 McDATA Parkway
Broomfield, Co. 80021
USA

Phone +1 720 558 4142
Fax +1 720 558 8999
EMail: mike.o'donnell@mcdata.com

Charles Monia

EMail: cmonia@pacbell.net

Milan J. Merhar
Sun Microsystems
43 Nagog Park
Acton, MA 01720
USA

Phone: +1 978 206 9124
EMail: milan.merhar@sun.com

Full Copyright Statement

Acknowledgement

Funding for the RFC Editor function is currently provided by the Internet Society.